Early warning of slight changes in systems and plants with application to condition based maintenance

Qinghua Zhang, Michèle Basseville, Albert Benveniste

UNITÉ DE RECHERCHE
INRIA-RENNES

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

1 9 9 2

25ème anniversaire

N° 1750

Programme 5
Traitement du Signal,
Automatique et Productique

EARLY WARNING OF SLIGHT CHANGES IN SYSTEMS AND PLANTS WITH APPLICATION TO CONDITION BASED MAINTENANCE

Qinghua ZHANG
Michèle BASSEVILLE
Albert BENVENISTE

Septembre 1992

RR-1750

IRISA

INSTITUT DE RECHERCHE EN INFORMATIQUE
ET SYSTEMES ALEATOIRES

Campus Universitaire de Beaulieu
35042 - RENNES CEDEX FRANCE
Tél. : 99 84 71 00 - Télex : UNIRISA 950 473 F
Télécopie : 99 38 38 32

Early warning of slight changes in systems and plants with application to condition based maintenance

Qinghua ZHANG*, Michèle BASSEVILLE, Albert BENVENISTE †

July 17, 1992
Publication Interne n° 671 - Juillet 1992 - 32 pages - Programme 5

Abstract

Techniques for early warning of slight changes in systems and plants are useful for condition based maintenance. In this paper we present an approach for this problem. This approach is based on the so-called "asymptotic local" approach for change detection previously introduced by the same authors. Its original principle consists in characterizing a system via some identified model, and then to monitor its changes using some data-to-model distance also derived from identification techniques. We show here that this method is of much wider applicability : model reduction can be enforced, biased identification procedures can be used, and finally one can even get rid of identification and use instead some much simpler Monte-Carlo estimation technique prior to change detection. Experiments on AR model are reported and an example from gas turbine industry is briefly discussed.

Keywords: failure detection, identification.

*during 1992, with Linköping Tekniska Högskola, Institutionen för Systemteknik, 581 83 Linköping, Sweden.

†This work is supported by Alcatel Alsthom Recherche and European Gas Turbine SA ; Sophie Jousset and Bruno Clavieras, both from Alcatel Alsthom Recherche, and Thierry Pourchot from European Gas Turbine SA, are gratefully acknowledged for this cooperation.

Détection Précoce de Changements Petits dans les Procédés Industriels - Application à la Maintenance Préventive

17 juillet 1992

Résumé

Les techniques permettant la détection précoce de changements petits dans les systèmes et procédés industriels sont utiles pour la maintenance préventive. Dans cet article nous proposons une approche possible pour résoudre ce problème. Cette approche est basée sur la méthode dite "asymptotique locale" pour la détection de changements (ruptures) que nous avons introduite précédemment. Son principe original consiste à caractériser un système à l'aide d'un modèle que l'on identifie sur des données, et ensuite à surveiller des déviations par rapport à ce modèle, en utilisant une sorte de distance entre données et modèle également déduite de techniques d'identification. Nous montrons ici que cette méthode peut en fait être appliquée de manière beaucoup plus générale : on peut surveiller avec un modèle *réduit*, on peut utiliser une identification *biaisée*, et on peut même se passer de l'identification et utiliser, pour la caractérisation du système sain, une méthode d'estimation de type Monte-Carlo beaucoup plus simple. On montre des résultats concernant d'une part un modèle AR scalaire, d'autre part la surveillance de la chambre de combustion d'une turbine à gaz.

Mots-clés: détection de pannes, identification.

1 Introduction

Applications of change detection techniques typically fall into two broad categories [2] :

- Quick detection of failure in sensors, actuators, or other components of a control system [2, 11, 14] ; this area is commonly referred to as that of **failure detection and identification.**
- Early detection of slight changes in the behavior of a plant and its interpretation and diagnostics ; although extremely important in the industry, this area has been much less investigated. It is usually referred to as **"condition-based" or "predictive" maintenance** [3].

This paper is mainly devoted to the second aspect. We use the statistical techniques for condition-based maintenance which we introduced in [6, 7] where the so-called **"local approach"** has been proposed as a general method for this problem. The purpose of the present paper is to investigate the robustness properties of this method with respect to issues such as use of a biased identification procedure, or use of model sets which do not contain the system in consideration.

Discussing a simple example

Let us first outline our approach on a very elementary example : the case of a scalar linear regression of the form

$$y_k = \varphi_k^T \theta + v_k \tag{1}$$

where y_k and the *exogenous* variables φ_k are observed, v_k is some white noise independent of φ_k, and the parameter vector $\theta \in \mathbf{R}^{d'}$ characterizes the system. We perfectly know how to identify θ based on measurements. The recursive least squares algorithm

$$\left\{ \begin{array}{rcl} \theta_k & = & \theta_{k-1} + R_k^{-1} \varphi_k \left(y_k - \varphi_k^T \theta_{k-1} \right) \\ R_k & = & R_{k-1} + \varphi_k \varphi_k^T \end{array} \right. \tag{2}$$

or simply the LMS algorithm

$$\theta_k = \theta_{k-1} + \gamma_k \varphi_k \left(y_k - \varphi_k^T \theta_{k-1} \right) \tag{3}$$

where γ_k is some suitable scalar gain, can be used for this purpose.

Now assume that we have a **nominal model** for such a system, which we denote by θ_o. Such a nominal model may be known either from prior physical knowledge of the system or from some prior identification based on a set of training data. We are interested in detecting (and possibly diagnosing) slight deviations of the actual system from its nominal behavior θ_o. Different techniques have been proposed in the literature for this purpose [2], they generally re-use the above identification procedures and track significant changes in the parameter θ_k

estimated on-line. In other words, changes are sought for by inspecting a suitable distance between the nominal model θ_o and its actual estimate θ_k.

In contrast, our approach consists in checking whether the current measurements still are in agreement with the nominal model θ_o, i.e., we consider a signal-to-model distance instead. In the present case of a linear regression, this can be done as follows. Both identification algorithms (2) and (3) involve the vector-valued function

$$H(\theta; y_k, \varphi_k) \triangleq \varphi_k \left(y_k - \varphi_k^T \theta\right) \tag{4}$$

and it turns out that deviations from the nominal model θ_o can be simply detected by estimating, for $\theta = \theta_0$, the average

$$h(\theta) \triangleq \mathbf{E} H(\theta; y_k, \varphi_k) \tag{5}$$

where expectation is taken jointly over regression vector and observation. The fact that θ_o is still the actual true system is characterized by the condition

$$h(\theta_o) = 0$$

which also expresses that θ_o is an equilibrium of the ODE associated with the recursive algorithms (2) or (3), see [7]. So it is expected that $H(\theta; y_k, \varphi_k)$ introduced in (4) is a convenient tool for early warning of deviations from nominal model. How this can be achieved is explained next.

We first make the actual (and unknown) true system explicit in (5), we denote it by θ_* and rewrite $h(\theta)$ as

$$h(\theta_*, \theta) \triangleq \mathbf{E}_{\theta_*} H(\theta; y_k, \varphi_k) \tag{6}$$

where the notation $\mathbf{E}_{\theta_*}$ denotes expectation taken jointly over regression vector and observation, when the true system is θ_*. Using the particular form (1) of the linear regression, we can rewrite this explicitly as follows :

$$\begin{aligned} H(\theta; y_k, \varphi_k) &= \varphi_k \left(\varphi_k^T \theta_* - \varphi_k^T \theta + v_k\right) \\ h(\theta_*, \theta) &= \mathbf{E}\left(\varphi_k \left(\varphi_k^T \theta_* - \varphi_k^T \theta + v_k\right)\right) \end{aligned} \tag{7}$$

where expectation is now taken over noise and (exogenous) regression vector, which do not depend on θ_*. It is clear on (7) that h satisfies the condition

$$h(\theta, \theta) \equiv 0 \quad \forall \theta \tag{8}$$

which is a necessary condition for both algorithms (2) and (3) to identify the true system [7]. Next assume we have a set of samples of new measurements $y_1, \cdots, y_N$ and $\varphi_1, \ldots \varphi_N$ taken from the true system θ_*, and consider the cumulative sum

$$\begin{aligned} D_N(\theta_o) &\triangleq \frac{1}{\sqrt{N}} \sum_{k=1}^{N} H(\theta_o; y_k, \varphi_k) \\ &= \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \varphi_k \varphi_k^T (\theta_* - \theta_o) + \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \varphi_k v_k \\ &\triangleq (a) + (b) \end{aligned}$$

Since both φ_k and v_k are uncorrelated and zero mean, $\varphi_k v_k$ is also zero mean and, by Central Limit Theorem, (b) converges to some zero mean Gaussian distribution which do not depend on the unknown true system θ_* (because the φ_k are exogenous). The dependency of $D_N(\theta_o)$ on the true system is concentrated on (a). To handle it in the proper asymptotic way, the **local** approach consists in normalizing the deviation $\theta_* - \theta_o$ as follows

$$\theta_* - \theta_o = \frac{\tilde{\theta}}{\sqrt{N}} \tag{9}$$

where $\tilde{\theta}$ is now considered as fixed. Hence (a) rewrites

$$(a) = \left(\frac{1}{N}\sum_{k=1}^{N}\varphi_k\varphi_k^T\right)\tilde{\theta} \to (\mathrm{Cov}\varphi)\,\tilde{\theta}$$

where $\mathrm{Cov}\varphi$ denotes the covariance matrix of φ. To summarize, if (9) is assumed, we have

$$D_N(\theta_o) \to \mathcal{N}\left(\mathrm{Cov}\varphi \cdot \tilde{\theta}, R\right) \tag{10}$$

by Central Limit Theorem, where R is the asymptotic covariance matrix of (b). The following facts are of importance in (10), they can be generalized as we shall see later :

1. $D_N(\theta_o)$ is asymptotically Gaussian for N large and for small deviations from the nominal model.

2. Small deviations from the nominal model are entirely reflected in the mean of $D_N(\theta_o)$, not in its covariance matrix, i.e., we have

$$D_N(\theta_o) \simeq \mathcal{N}\left(F \cdot \tilde{\theta}, R\right) \tag{11}$$

3. The matrices F and R in (11) are easily estimated even for general identification problems since they depend only on the *available* nominal model θ_o, so Monte-Carlo estimates can be at least relied on when no closed form formula is known.

Point 3. holds true also for complex nonlinear systems as we shall see later, and this is in contrast with asymptotic error covariances of identification procedures, which are difficult to estimate except for linear systems.

The conclusion of this discussion is that *small changes in θ_* are reflected in changes in the mean of $H(\theta; y_k, \varphi_k)$ which, moreover, can be processed as if it was a sequence of independent Gaussian random variables with covariance R.*

Drawing our general approach from the simple example

Based on the previous example, we can draw our *general method* for detecting small deviations from some nominal model in a system $\mathcal{L}_\theta$:

1. Select some recursive identification procedure suitable to identify the true system in the class $(\mathcal{L}_\theta)$, say
$$\theta_k = \theta_{k-1} + \gamma_k H(\theta_{k-1}, X_k) \tag{12}$$

2. Consider
$$Z_k = H(\theta_o, X_k)$$
where θ_o is some nominal model, and (X_k) are new measurements taken from the system.

3. Monitor (Z_k) using standard techniques for on-line or off-line detection of changes in the mean of an independent Gaussian process [6, 2, 5]. Various refinements can be performed at this point to diagnose the origin of the changes.

In our discussion of the linear regression example, the following conditions have played a key role :

The true system is in the model set : we have assumed that (y_k, φ_k) were really taken from a regression system of the form (1).

The monitoring procedure was derived from a bias-free identification algorithm : this was used in (8) and the subsequent discussion.

The same identification procedure was used, both to derive the monitoring procedure, and to estimate θ_o when this nominal model is indeed fitted on some training data set.

The purpose of the present paper is to show that these requirements which were first presented in [6, 7] can be drastically relaxed : the true system does not need to belong to the model set, biased identification algorithms can be used to design the tests, and we can even avoid using system identification procedures to characterize the nominal behavior. This will be illustrated on two examples, the case of AR processes, and an example of non-linear plant, namely the monitoring of the set {combustion chamber + turbine} of a gas turbine power generator.

Section 2 is the core of the paper. In subsection 2.1 we introduce the fairly wide class of *systems* to which our local methods applies. The *models* that can be used for monitoring are discussed in subsection 2.2. Subsection 2.3 is devoted to the central limit theorem supporting our approach, and a first method for monitoring is presented which relies upon identification of the nominal model. In subsection 2.4, we show how nominal model identification can be avoided and a second method for monitoring is presented and issues of performance are discussed. The practical implementation of the approach is presented in subsection 2.5. Section 3 is devoted to reports about experiments. In subsection 3.1, we report on AR process monitoring with model reduction, using the second method. Finally, the industrial application which called for the generalization of our original method is briefly presented in subsection 3.2.

2 The local approach

In this section we present the theoretical background of our method, following [6, 7]. However we shall deviate from these references in order to relax the requirements we have discussed in the previous section. This will considerably enlarge the range of validity of this approach.

2.1 The considered systems

They are of the form

$$\begin{aligned} \mathbf{P}\left(\xi_k \in A \,|\xi_{k-1}, \xi_{k-2}, \cdots\right) &= \int_A \pi_\Theta\left(\xi_{k-1}, dx\right) \\ y_k &= f\left(\xi_k\right) \end{aligned} \tag{13}$$

where

- Θ is some vector parameter belonging to some open domain $\mathcal{D}$ in $\mathbf{R}^d$ and f some function ;
- y_k are the observations ;
- ξ_k is not observed in general ; for Θ fixed in $\mathcal{D}$, it is a Markov chain with transition probability $\pi_\Theta(\xi, dx)$ having a unique invariant probability. We denote by $\mathbf{P}_\Theta$ the corresponding stationary law for the Markov chain (ξ_k), and by $\mathbf{E}_\Theta$ the corresponding expectation[1].

Model (13) is extremely general [7]. It encompasses all usual linear models (AR, ARMA(X), state space) and also general non-linear dynamical systems having some Markov extended state. For instance, if (y_k) is an ARMA (p, q) process with innovation ν_k, take for Θ the collection of ARMA parameters, and for state $\xi_k^T = (y_k, \cdots, y_{k-p}; \nu_k, \cdots, \nu_{k-q})$ so that function f in (13) just consists in selecting the first component of ξ. For the regression example we discussed in the introduction, and assuming all involved random sequences to be i.i.d., simply take for the observations the vector $\left(y_k, \varphi_k^T\right)$, take also $\Theta = \theta$ (i.e., model $\equiv$ system), and finally take also $\left(y_k, \varphi_k^T\right)$ as being the state occurring in the first equation of (13).

2.2 The model and basic statistics

We do not require that equations (13) involving Θ are used as a model to monitor the system. We rather assume that we use some alternative model with a generic parameter θ belonging to some open domain D in $\mathbf{R}^{d'}$ and with a structure possibly different from that of (13)[2].

[1] The reader may simply think of $\mathbf{E}_\Theta f(\xi_\cdot)$ as being the $\lim_{k\to\infty} \mathbf{E}_\Theta\, f(\xi_k)$, and $\mathcal{D}$ as the domain of stability of the considered Markov chain.

[2] This was not the case in the regression example, but this is precisely one of the conditions of this example we want to relax ; for instance, we may want to enforce model reduction.

Our key tool will be the notion of a **basic statistics**:

$$H(\theta, X_k) \tag{14}$$

where X_k is an auxiliary sequence defined below. In selecting such a statistics we require the following conditions :

Condition 1 (generation of X_k) *For θ fixed in D, (X_k) is the output of some known dynamical system of the form*

$$\begin{cases} \Xi_k &= G(\theta, \Xi_{k-1}, y_k) \\ X_k &= g(\Xi_k) \end{cases} \tag{15}$$

For the case of the regression example, we simply have $X_k^T = \left(y_k, \varphi_k^T\right)$, i.e., the X auxiliary sequence is just the observation itself. The form (15) for generating X_k is however useful for more complex problems such as the monitoring of ARMA systems.

Condition 2 (averaging) *For any pair (Θ, θ) in $\mathcal{D} \times D$, the following expectations exist*

$$h(\Theta, \theta) \triangleq \mathbf{E}_\Theta H(\theta, X_k) \tag{16}$$

$$R(\Theta, \theta) \triangleq \sum_{k=-\infty}^{+\infty} \mathrm{cov}_\Theta \left[H(\theta, X_0), H(\theta, X_k)\right] \tag{17}$$

where cov_Θ denotes covariance under law $\mathbf{P}_\Theta$ and X_k is generated via (15) using parameter θ. In addition, we require the matrix $R(\Theta, \theta)$ to be positive definite.

The corresponding objects have also been introduced for the regression example, note however that R does not depend upon (θ, Θ) in this simple case, cf. Equation (10).

One of the key properties in establishing our method for the case of linear regression is the identity (8). It should be clear that such an identity cannot hold any more when model reduction is enforced (i.e., when Θ-space and θ-space are different), or when the basic statistics $H(\theta, X)$ is associated with some biased identification algorithm. Therefore it is desirable to weaken the requirement that identity (8) should hold.

Condition 3 (bias function) *The function $h(\Theta, \theta)$ is continuously differentiable, and the partial derivative $\frac{\partial}{\partial \theta} h(\Theta, \theta)$ is a square invertible matrix. Consequently, by the implicit function theorem (see for instance theorem (10.2.1) in [9]), the equation $h(\Theta, \theta) = 0$ defines a unique map $\mathcal{F}$, continuously differentiable from the Θ-space $\mathcal{D}$ into the θ-space D, such that*

$$\theta = \mathcal{F}(\Theta) \quad \textit{iff} \quad h(\Theta, \theta) = 0 \tag{18}$$

$\mathcal{F}$ is called the **bias function** *of the basic statistics H.*

Comment. We call $\mathcal{F}$ a *bias function*, because the basic statistics $H(\theta, X_k)$ (and therefore its expectation $h(\Theta, \theta)$) is typically associated with some recursive identification procedure and $\mathcal{F}$ specifies the relationship between the system parameter Θ and the identified model parameter θ for this procedure. The condition that such a bias function exists, which associates a *unique* θ-model with a given Θ-system expresses that we allow model reduction ($\mathcal{F}$ does not need to be one-to-one) but we forbid over parameterization (where several θ's can be associated with a given Θ via the equation $h(\Theta, \theta) = 0$). Note that if the nominal model under consideration was actually identified using *this* identification procedure, then the true system before change and the nominal model are related via $\mathcal{F}$. There are situations, however, where one cannot use the identification procedure associated with the basic statistics for identifying the nominal model, we shall discuss how such a problem can be handled in Subsection 2.4.

Summarizing the different objects and their status. The following objects are known and available to the user :

- the observations y_k,
- the adjustable parameter θ and the dynamical system (15),
- the basic statistics $H(\theta, X_k)$ (where X_k is generated from observations y_k via (15)) and the nominal model θ_o.

On the other hand, the following objects are not accessible to the user, they are only mathematical objects satisfying some conditions :

- the true system Θ and law $\mathbf{P}_\Theta$,
- the averaged quantities $h(\Theta, \theta)$ and $R(\Theta, \theta)$,
- the bias function $\mathcal{F}$.

The following can be said about these unknown objects:

- the true system Θ and law $\mathbf{P}_\Theta$ are unknown, but Monte-Carlo estimates can be obtained using sample observations from this distribution;
- the averaged quantities $h(\Theta, \theta)$ and $R(\Theta, \theta)$ are unknown but can be estimated via the above mentioned technique (since θ in turn is known);
- the bias function $\mathcal{F}$ is unknown and is not at all needed for implementing the algorithm, it is just a tool for the mathematical analysis.

Example : Regression. In the linear regression example, Θ and θ correspond to the same model structure, observations are the pair (y_k, φ_k), we have $X_k^T = \left(y_k, \varphi_k^T\right)$, and $\mathcal{F}$ is just identity.

In our more general situation, we may have systems not in the model set (so Θ is different from θ). For a system Θ, the "true model" is characterized by (18), i.e., via the bias function $\theta = \mathcal{F}(\Theta)$. Let us look at another example.

Example : ELS. If (y_k) is ARMA(p,q) with innovation (ε_k), we take for Θ the ARMA parameters and the domain $\mathcal{D}$ is the domain of stable and minimum phase Θ–models. Now assume that an ARMA(p',q') is used for monitoring, with $p' \leq p$ and $q' \leq q$. Collect the corresponding ARMA(p',q') parameters into a vector θ and call D the domain of minimum phase θ–models. For θ fixed in D, denote by $e_k(\theta)$ the prediction error associated with θ, and set:

$$\varphi_k^T(\theta) = (\ y_{k-1},\ \ldots,\ y_{k-p';}\ e_{k-1}(\theta),\ \ldots,\ e_{k-q'}(\theta)\ )$$

and:

$$X_k^T(\theta) = (\ e_k(\theta);\ \varphi_k^T(\theta)\ )$$

which can be shown [7] to be generated via a model of the form (15). Then:

$$H(\theta, X_k) \triangleq \varphi_k(\theta) e_k(\theta)$$

is the basic statistics associated with the ELS algorithm for ARMA(p',q') model fitting. The bias function $\mathcal{F}$ associates with the system Θ the reduced order model θ such that $\mathbf{E}_\Theta \varphi_k(\theta) e_k(\theta) = 0$.

2.3 Central limit theorems and the first implementation of the local approach

We shall now present the limit theorems which justify our approach. We shall distinguish between off-line detection problems which we shall refer to as **model validation** problems, and on-line detection problems which we shall call **change detection** problems. Corresponding mathematical problem statements are given now. In both cases, some nominal model θ_o of the system is available (we assume in this subsection that θ_o is identified using some training data from the system under consideration), and we denote by Θ_* the (unknown) true system, these notations will always be used in the sequel. Now we are making a record of new measurements $y_1, \cdots, y_N$, and the detection problems are stated as follows:

Problem 1 (model validation) *test $\mathcal{H}_1$ against $\mathcal{H}_0$, where*

$$\mathcal{H}_0 \ :\ \mathcal{F}(\Theta_*) = \theta_o \ \text{ for } \ k = 1, \cdots, N$$

$$\mathcal{H}_1 \ :\ \mathcal{F}(\Theta_*) = \theta_o + \frac{\tilde{\theta}}{\sqrt{N}} \ \text{ for } \ k = 1, \cdots, N$$

where $\tilde{\theta}$ is an unknown but fixed hypothesized change and $\mathcal{F}$ is the bias function of the considered basic statistics H.

Problem 2 (on-line change detection) *test $\mathcal{H}_1$ against $\mathcal{H}_0$, where*

$$\begin{aligned} \mathcal{H}_0 &: \mathcal{F}(\Theta_*) = \theta_o \text{ for } k = 1, \cdots, N \\ \mathcal{H}_1 &: \text{there exists some unknown } \tau \in (0,1) \text{ such that} \end{aligned}$$

$$\begin{aligned} \mathcal{F}(\Theta_*) &= \theta_o \text{ for } 1 \leq k < \tau N \\ \mathcal{F}(\Theta_*) &= \theta_o + \frac{\tilde{\theta}}{\sqrt{N}} \text{ for } \tau N \leq k \leq N \end{aligned}$$

where $\tilde{\theta}$ and $\mathcal{F}$ are as above.

Remark that if θ_o is actually identified by an identification procedure based on $H(\theta, X_k)$ associated with $\mathcal{F}$, then for the system Θ_* before change, we have $\mathcal{F}(\Theta_*) = \theta_o$ by definition of $\mathcal{F}$. Therefore problem 1 means to test whether θ_o is in agreement with the whole new record $y_1, \cdots, y_N$, or in other words, whether the true system Θ_* (or more precisely, $\mathcal{F}(\Theta_*)$) has changed since the identification of θ_o. This is what is often performed in condition based maintenance when some test data are collected at different periods to check for degradations. Problem 2 considers permanent monitoring and is concerned with detecting the change as soon as possible. The second problem will require limit theorems for stochastic processes, so we defer its theoretical investigation to Appendix A. In contrast, problem 1 will rely on a simple Central Limit Theorem so we investigate it in the core of the paper. We should emphasize that the way in which the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ are formulated in both cases expresses that *only changes in $\mathcal{F}(\Theta)$ are going to be detected.* In other words, changes in Θ that are not reflected by the bias function $\mathcal{F}$ cannot be detected.

Introduce the following cumulative sum (*Cusum*)

$$D_N(\theta) \triangleq \frac{1}{\sqrt{N}} \sum_{k=1}^{N} H(\theta, X_k) \tag{19}$$

We refer the reader to the appendix of [6] for some additional technical mixing conditions for the following theorem to hold :

Theorem 1 *Under Conditions 1, 2, 3, and the hypothesis*

$$\mathcal{F}(\Theta_*) = \theta_o + \frac{\tilde{\theta}}{\sqrt{N}} \text{ for } k = 1, \cdots, N \ , \tag{20}$$

when $N \to \infty$, the Cusum $D_N(\theta_o)$ converges weakly to the Gaussian distribution

$$\mathcal{N}\left(-\frac{\partial}{\partial \theta} h(\Theta_*, \mathcal{F}(\Theta_*)) \cdot \tilde{\theta} \ ; \ R(\Theta_*, \mathcal{F}(\Theta_*)) \right) \tag{21}$$

where $h(\Theta, \theta)$ and $R(\Theta, \theta)$ are defined in (16,17).

Proof : we follow closely the proof in the appendix of [6]. Using (16) and following exactly the argument of [6], we get

$$R(\Theta_*, \theta_o)^{-1/2} \left(D_N(\theta_o) - \sqrt{N}\, h(\Theta_*, \theta_o) \right) \to \mathcal{N}(0, I)$$

where Θ_* is as in (20). Performing a first order Taylor expansion and using (18) and (20), we get

$$\begin{aligned} \sqrt{N} h(\Theta_*, \theta_o) &\simeq \sqrt{N} h(\Theta_*, \mathcal{F}(\Theta_*)) - \frac{\partial}{\partial \theta} h(\Theta_*, \mathcal{F}(\Theta_*)) \cdot \tilde{\theta} \\ &= - \frac{\partial}{\partial \theta} h(\Theta_*, \mathcal{F}(\Theta_*)) \cdot \tilde{\theta} \end{aligned}$$

which proves the theorem.□

Notations. For the sake of convenience, we shall write in the sequel

$$h(\Theta_*) \quad \text{instead of} \quad h(\Theta_*, \mathcal{F}(\Theta_*)) \tag{22}$$

$$h'(\Theta_*) \quad \text{instead of} \quad \frac{\partial}{\partial \theta} h(\Theta_*, \mathcal{F}(\Theta_*)) \tag{23}$$

$$R(\Theta_*) \quad \text{instead of} \quad R(\Theta_*, \mathcal{F}(\Theta_*)) \tag{24}$$

Comment. The statement of Theorem 1 is slightly informal : the distribution (21) is *not* fixed since it depends on the true system Θ_*, which in turn depends on the length N of the sample as indicated in (20). So we should have written instead that "Cusum $D_N(\theta_o)$ and Distribution (21) are *contiguous*" [12], which means that, when N tends to infinity, for any positive and continuous function f, $\mathbf{E}f[D_N(\theta_o)] - \mathbf{E}f[Z_N]$ tends to zero if Z_N behaves according to distribution (21). In fact, when the bias function $\mathcal{F}$ is identity, then we can replace in formula (21) the true system Θ_* (which depends on N) by the *fixed* nominal model θ_o, so the statement of the theorem is correct in this case ; this is exactly the way the corresponding Central Limit Theorem is stated in [6], see also [7] [5].

Theorem 1 is the basis for deriving off-line monitoring procedures. The corresponding theorem for on-line procedures is proposed in Appendix A. These justify that we can transform the detection of changes in system Θ_* into the asymptotically equivalent one: the detection of changes in the mean of the basic statistics, as if it was an independent Gaussian process. The reader is referred to [6, 2, 5] for the design of off-line and on-line algorithms for the detection of changes in the mean of an independent Gaussian process. Such an algorithm is given in the application section of this paper.

Summary. What we have shown at this point is the following :

1. *Model reduction* can be performed when monitoring is applied to the system (using θ instead of Θ). Changes on the true system Θ that can be detected are actually those which are visible on the reduced model θ.

2. Basic statistics $H(\theta, X_k)$ can be considered to be associated with some *biased identification* procedure. Then this identification procedure is used to estimate the nominal model θ_o based on a sample of training data, and $H(\theta_o, X_k)$ is used for subsequent (on-line or off-line) monitoring of the system.

The corresponding general method for detecting slight changes is recalled here :

Method 1 (biased identification of nominal model)

1. *Select some recursive identification procedure suitable to identify the "best" nominal model, say*
$$\theta_k = \theta_{k-1} + \gamma_k H(\theta_{k-1}, X_k)$$
and denote by θ_o this identified nominal model. It is in general related to the true system via some nontrivial bias function $\mathcal{F}$ which can involve model reduction.

2. *Consider*
$$Z_k = H(\theta_o, X_k)$$
where (X_k) are generated using new measurements taken from the system.

3. *Monitor (Z_k) using standard techniques for on-line or off-line detection of changes in the mean of an independent Gaussian process [6, 2, 5]. Various refinements can be performed at this point to diagnose the origin of the changes.*

2.4 Replacing identification by Monte-Carlo estimation and the second implementation of the local approach

In the application example of Section 3.2 we are faced with the following problem: we were not able to develop a conveniently stable and convergent recursive algorithm for identifying the nominal model θ_o based on observations from the nominal system. We have instead to rely on some off-line identification procedure from which no basic statistics can be easily derived. So, what we do is the following:

1. we design a well conditioned (but biased) off-line procedure for identification, this yields a nominal model θ_o,

2. on the other hand we consider the basic statistics associated with a standard prediction error method arising from some least squares criterion (see (41,42) in Section 3.2). Note that this standard prediction error method cannot be used for identifying our system because of its bad convergence behavior on this application.

This yields some basic statistics (see (45) in Section 3.2)) we denote by $H(\theta_o, X_k)$. Denote by $\mathcal{F}$ its bias function. Then, since our nominal model θ_o was not identified using H, *we do not have $\mathcal{F}(\Theta_*) = \theta_o$ before any change of Θ_**, and this does not conform to Hypothesis $\mathcal{H}_0$ in the statement of Problems 1 and 2. Therefore Theorem 1 does not apply and Method 1 cannot be used. The alternative method we propose is the following :

Method 2 (Monte-Carlo estimation of the nominal bias)

1. *Let the true system be generically denoted by Θ and the actual model we use by θ.*

2. *Choose a nominal model θ_o (not necessarily by identification).*

3. *Estimate on the training data the bias*

$$h_o \triangleq \mathbf{E}_{\Theta_*} H(\theta_o, X_k) \tag{25}$$

and the variance

$$R_o \triangleq \sum_{k=-\infty}^{+\infty} \mathrm{cov}_{\Theta_*}\left[H(\theta_o, X_0), H(\theta_o, X_k)\right] \tag{26}$$

This step replaces nominal model identification, and is performed via Monte-Carlo estimation:

$$h_o = \mathbf{E}_{\Theta_*} H(\theta_o, X_k) \approx \hat{h}_o \triangleq \frac{1}{K}\sum_{k=1}^{K} H(\theta_o, X_k) \tag{27}$$

where $\{X_k : k = 1, \ldots . K\}$ are generated via (15) using $\theta = \theta_o$, with observations taken before any change of Θ_*, and similarly for R_o.

4. *Use $H_o(\theta_o, X_k) \triangleq H(\theta_o, X_k) - h_o$ as a basic statistics, and proceed as for the standard case.*

To justify Method 2, simply note that H_o satisfies Conditions 1, 2, 3, if H does. Furthermore, the bias function $\mathcal{F}_o$ associated with H_o is defined by $h(\Theta, \theta) = h_o$, so that $\theta_o = \mathcal{F}_o(\Theta_*)$ holds for system Θ_* before change. Hence, Theorem 1 applies to $H_o(\theta_o, X_k)$ if we replace in the statements of problems 1 and 2 $\mathcal{F}$ by $\mathcal{F}_o$.

Discussion : issues of performance

The nominal model θ_o used in method 2 is not necessarily identified and can be, at least in principle, arbitrarily selected. The intuition behind this procedure is that Monte-Carlo estimation of the bias h_o can replace nominal model identification. Clearly, Method 2 is much easier to implement than Method 1, since Monte-Carlo estimation of the bias at some chosen nominal model is easier than identifying the "true" nominal model corresponding to the nominal system. So, what is the interest of Method 1 ? In fact, experiments reported in Section 3 show that a loss of efficiency results from selecting in Method 2 a nominal model far from the "true" one. So let us discuss this issue of efficiency.

It is shown in [4], formula (21), that the efficacy of testing $\mu \neq 0$ against $\mu = 0$ in $\mathcal{N}(M\mu, R)$ is measured by the quantity

$$\mathrm{Eff} \triangleq \mathrm{Tr}\, MR^{-1}M^T$$

In our case, this yields

$$\text{Eff}_{\Theta_*}(H) = \text{Tr}\ \Delta^{-1}, \quad \text{where}\ \ \Delta \triangleq h'^{-1}(\Theta_*)\, R(\Theta_*)\, h'^{-T}(\Theta_*) \tag{28}$$

On the other hand, it is shown in [7], formula (3.2.10) of Chapter 3, that Δ introduced in (28) is the *intrinsic quality criterion* associated with the basic statistics $H(\theta, X_k)$. This criterion is used in [7] to assess algorithm (12) with regard to both its tracking capability under nonstationary environment and convergence rate under steady environment. Larger values of Δ correspond to better quality of H. It is also shown that maximal Δ is the inverse Fisher matrix, and is achieved by taking for H the *efficient score* [5], i.e., minus the gradient of the log-likelihood function. As a conclusion, we can assert that *the most efficient implementation of the local approach is achieved by applying Method 1 and selecting a basic statistics associated with a recursive maximum likelihood identification procedure.*

We can be more specific in assessing the loss of performance resulting from a poor choice of the nominal model in Method 2 for the simple case of linear regression we discussed in the introduction.. So, we consider again the regression example of the introduction, assuming that all involved sequences of random variables are i.i.d.. So, consider again the basic statistics

$$H(\theta; y_k, \varphi_k) \triangleq \varphi_k \left(y_k - \varphi_k^T \theta\right)$$

Select some nominal model θ_o and denote by h_o the mean of H at this nominal model when the true system is actually θ_*. Thus the modified basic statistics we have to consider for Method 2 is

$$H_o(\theta_o; y_k, \varphi_k) \triangleq H(\theta_o; y_k, \varphi_k) - h_o$$

Let us compute criterion Δ for the so defined statistics H_o. We denote by h'_o, R_o the objects involved in the formula (28) for H_o. We have

$$h'_o \equiv h'$$

so that we only have to compare R and R_o. Denote by $\Sigma \triangleq \mathbf{E}_{\theta_*} \varphi\varphi^T$ the covariance matrix of the regressor, by $\tilde{\theta} = \theta - \theta_*$ the parameter error, and by σ^2 the variance of the noise v_k. Using these notations, the general formula (17) for R at any model θ yields in the case of the linear regression

$$\begin{aligned}
&R(\theta_*, \theta) \ = \\
&\sum_{k=-\infty}^{+\infty} \mathbf{E}_{\theta_*} \left[\left(\varphi_k\varphi_k^T - \Sigma\right)\tilde{\theta} + \varphi_k v_k\right] \left[\left(\varphi_0\varphi_0^T - \Sigma\right)\tilde{\theta} + \varphi_0 v_0\right]^T \\
&= \mathbf{E}_{\theta_*} \left[\left(\varphi_0\varphi_0^T - \Sigma\right)\tilde{\theta} + \varphi_0 v_0\right] \left[\left(\varphi_0\varphi_0^T - \Sigma\right)\tilde{\theta} + \varphi_0 v_0\right]^T \\
&= \sigma^2\Sigma + \mathbf{E}_{\theta_*} \left(\varphi\varphi^T - \Sigma\right)\tilde{\theta}\tilde{\theta}^T \left(\varphi\varphi^T - \Sigma\right) \\
&\triangleq \sigma^2\Sigma + \delta R(\tilde{\theta})
\end{aligned}$$

The loss of performance is visible in these equalities, since the additional matrix $\delta R(\tilde{\theta})$ is positive and is larger for larger $\tilde{\theta}$'s.

2.5 Practical use of the local approach for detection and diagnosis

So far our discussion was based on the local approach, i.e., it involved considering alternative hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ that get closer when the sample length tends to infinity (recall that the hypothesized change is of the form $\tilde{\theta}/\sqrt{N}$). On the other hand, actual changes that may occur in practice are of some size and the normalization via the factor of $\sqrt{N}$ is irrelevant in practice. So it is needed to discuss precisely how our theoretical investigation can result in some effective method that can be used in practice.

We concentrate again on Problem 1 and refer the reader to Appendix A for Problem 2. Introduce

$$Z_k = H(\theta_o, X_k)$$

and consider, on one hand, problem 1, and on the other hand, the following problem :

Problem 3 (Gaussian equivalent model validation problem) *Assume (Z_k) is an independent sequence of random variables, and test $\mathcal{K}_1$ against $\mathcal{K}_0$, where*

$$\begin{aligned} \mathcal{K}_0 &: Z_k \sim \mathcal{N}(0, R) \quad \text{for} \quad k = 1, \cdots, N \\ \mathcal{K}_1 &: Z_k \sim \mathcal{N}\left(-h' \cdot \frac{\tilde{\theta}}{\sqrt{N}}, R\right) \quad \text{for} \quad k = 1, \cdots, N. \end{aligned}$$

where $\tilde{\theta}$ is an unknown but fixed hypothesized change, and R and h' are shorthand for the quantities defined in (24) and (23) respectively.

Then theorem 1 expresses that both problems are asymptotically equivalent for N large if the statistics $H(\theta_o, X_k)$ only is used. The practical use of this remark is the following.

1. Consider the basic statistics $Z_k = H(\theta_o, X_k)$, and handle it as if it was a sequence of independent Gaussian variables with covariance matrix R.

2. Check for deviations of the mean of Z_k from its zero nominal value. This can be performed either off-line (for model validation) or on-line (for change detection).

At this point it is worth to discuss the question of diagnostics and available prior information. Prior information is generally available in the form of a set $\tilde{\mathcal{C}}$ of candidate values for the deviation $\tilde{\theta}$. Such $\tilde{\mathcal{C}}$ could summarize all possible failures or alternatively characterize failures of a particular type. This "failure set" $\tilde{\mathcal{C}}$ is selected a priori by the designer, as well as the threshold, for the detection to be performed. So it is important that, when $\tilde{\theta}$ ranges over $\tilde{\mathcal{C}}$, the resulting mean $\mathbf{E}_\Theta Z_k$ ranges over a set which is *independent* of the actual sample length N, the latter one being unknown a priori. Such a property is simply achieved by selecting a failure set $\tilde{\mathcal{C}}$ which is invariant via the multiplication by a factor of the form $\sqrt{N}$. Hence, we require the following condition to be satisfied by $\tilde{\mathcal{C}}$:

$$\tilde{\mathcal{C}} = \lambda \tilde{\mathcal{C}} \quad \forall \lambda \in \mathbb{R}_+ \tag{29}$$

i.e., $\tilde{\mathcal{C}}$ should be a *cone*. Hence, instead of problem 3 we shall consider in practice problems of the following form for the case of model validation :

Problem 4 (use of prior information) *Test* $\tilde{\mathcal{K}}_1$ *against* $\mathcal{K}_0$*, where*

$$\tilde{\mathcal{K}}_1 : Z_k \sim \mathcal{N}\left(-h' \cdot \tilde{\theta}, R\right), \textit{for } k = 1, \cdots, N, \tilde{\theta} \in \tilde{\mathcal{C}}$$

and $\tilde{\mathcal{C}}$ *is a known cone, while* $\mathcal{K}_0$ *is as in problem 3.*

This is a classical Gaussian multiple hypotheses testing problem, and we refer the reader to [6, 3] for its use in diagnostics. Finally, the fact that different model sets can be handled for the system (characterized by parameter Θ) and the model used for change detection (characterized by parameter θ), can be used for *physical diagnostics* as outlined next. We might in fact know approximately model (13) from the designer (complex models are often used for simulation and design). In such a case, it is possible to formulate prior knowledge about possible failures in the "physical" Θ–space, and $\tilde{\mathcal{C}}$ in Problem 4 is then obtained via the Jacobian of the map from Θ–space into θ–space. Such an approach for diagnostics is discussed in [6] and extensively used in [3] for the case of vibration mechanics.

We are now ready to present and discuss the application example. All flexibility we have introduced in using our method will be used.

3 Numerical examples

In this section, we report on two different kinds of numerical examples. The first one is a simulation example, and is just intended to illustrate our purpose concerning model reduction and is dealing with AR process monitoring. The second one is a truly industrial application, and is at the origin of the present extensions of the previously existing local approach.

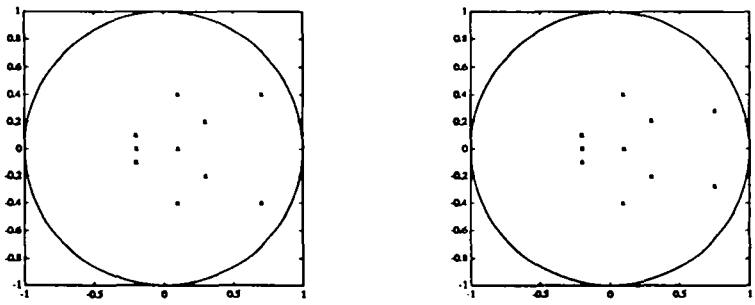
3.1 The case of an AR system with model reduction

To illustrate the new results of our local approach reported in the previous sections, we present in this subsection an simple example: the detection of change in an AR process. AR process is widely used for parametric signal analysis, see for instance [1, 10] for some application examples in speech signal segmentation and seismic signal processing. By modeling signal with AR process and detecting changes in the autoregressive parameters, our local approach for change detection constitutes a method for detecting signal spectrum changes. For the sake of completeness, we present here only results for the off-line approach. We shall present in the next subsection an application example of the on-line approach.

Problem statement

We use an AR process of order 10 to simulate signals:

$$y_k = \sum_{i=1}^{10} a_i y_{k-i} + v_k \tag{30}$$



Θ_o Θ_1

Figure 1: Poles of the AR process (30) with parameter Θ_o and Θ_1

where $a_1, \ldots, a_{10}$ are autoregressive parameters, y_k is assumed to be measured, and v_k is an i.i.d. noise with variance σ^2. We denote by Θ the autoregressive parameters:

$$\Theta = (a_1, \ldots, a_{10})^T \tag{31}$$

We denote by Θ_o the nominal parameter of the process with which signals before change are simulated, and by Θ_1 the parameter of the process after change. The values of Θ_o and Θ_1 we used are

$$\Theta_o = \begin{bmatrix} -1.700 & \times 10^{0} \\ 1.160 & \times 10^{0} \\ -2.980 & \times 10^{-1} \\ 1.520 & \times 10^{-2} \\ 3.212 & \times 10^{-2} \\ -7.986 & \times 10^{-3} \\ -9.942 & \times 10^{-4} \\ 8.737 & \times 10^{-4} \\ 7.105 & \times 10^{-5} \\ -1.437 & \times 10^{-5} \end{bmatrix} \qquad \Theta_1 = \begin{bmatrix} -1.785 & \times 10^{0} \\ 1.160 & \times 10^{0} \\ -2.980 & \times 10^{-1} \\ 1.520 & \times 10^{-2} \\ 3.212 & \times 10^{-2} \\ -7.986 & \times 10^{-3} \\ -9.942 & \times 10^{-4} \\ 8.737 & \times 10^{-4} \\ 7.105 & \times 10^{-5} \\ -1.437 & \times 10^{-5} \end{bmatrix}$$

for which only a change at the parameter a_1 is simulated. In both cases the variance of the noise is $\sigma^2 = 0.01$. The poles corresponding to the autoregressive parameters Θ_o and Θ_1 are depicted in figure 1.

To illustrate our local approach with model reduction, we shall use an AR model of order 2 to monitor the process of order 10. For this purpose the signal y_k is assumed to be modeled by an AR(2) process:

$$y_k = (y_{k-1} \;\; y_{k-2})\theta + w_k \tag{32}$$

with $\theta = (a_1, a_2)^T$ the autoregressive parameter and w_k the noise. We choose the basic statistics as in the introduction example, equation (4). Remark that the computation of the

basic statistics is related to model (32), not model (30) which is assumed to be unknown. The auxiliary variable X_k in this case is

$$X_k = (y_k, y_{k-1}, y_{k-2})^T$$

The training data is generated with Θ_o from which the nominal model will be identified, and the cumulative covariance matrix as well as the bias of the basic statistics will be estimated. Tests will be performed on signals generated with both Θ_o and Θ_1, and method 2 we presented in the previous section will be applied, in both the cases with an identified nominal model and with arbitrarily selected nominal models.

Estimation of the basic statistics bias and the cumulative covariance matrix

In order to apply method 2 we have to estimate the basic statistics bias h_o and the cumulative covariance matrix R_o as defined in equations (25)–(26).

Assume that $\{X_1, \ldots, X_K\}$ is generated from the training data $\{y_k\}$, θ_o is the chosen nominal model parameter, the bias h_o is estimated by:

$$\hat{h}_o = \frac{1}{K} \sum_{k=1}^{K} H(\theta_o, X_k), \tag{33}$$

then removing $\hat{h}_o$ from $H(\theta_o, X_k)$ yields

$$Z_k = H(\theta_o, X_k) - \hat{h}_o \tag{34}$$

According to the definition of R_o one can in principle use the following estimator of R_o:

$$\hat{R}_o \approx \frac{1}{n_0} \sum_{k=1}^{n_0} Z_k Z_k^T + \sum_{i=1}^{I} \frac{1}{n_i} \sum_{k=1}^{n_i} \left(Z_k Z_{k+i}^T + Z_{k+i} Z_k^T \right) \tag{35}$$

with some properly selected numbers I and $n_0, \ldots, n_I$.

The problem of this estimator of R_o is that it is not guaranteed to be positive definite, so instead we use the fact that R_o is also the covariance matrix of the cumulative sum $D_N(\theta_o)$ (as defined in (19)) for N large. Accordingly, we follow the suggestion of [8] and choose as an estimate of R_o :

$$\hat{R}_o = \frac{1}{L} \sum_{l=0}^{L-1} D_{N,l} D_{N,l}^T \text{ where :} \tag{36}$$

$$D_{N,l} = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} Z_{k+lN} \tag{37}$$

and $LN = K$ is the total length of the sample. Clearly $\hat{R}_o$ is generically positive definite for L large enough; on the other hand, N must be taken large enough so that D_N behaves approximately as a Gaussian random variable as stated in (21).

The local test

Assume that $\{X_1, \ldots, X_N\}$ are generated from some test data $\{y_k\}$, and Z_k is still defined by

$$Z_k = H(\theta_o, X_k) - \hat{h}_o$$

then according to the theoretical result of the previous section, detecting changes in the parameter Θ of process (30) is asymptotically equivalent to detecting changes in the mean of Z_k, as if it was an i.i.d. Gaussian variable with variance R_o. Therefore [6, 7] the test is computed as follows:

$$\begin{aligned} D_N &= N^{-\frac{1}{2}} \sum_{k=1}^{N} Z_k \\ S &= D_N^T \hat{R}_o^{-1} D_N \end{aligned}$$

The decision rule between hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ is

$$S \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda.$$

with a chosen threshold λ.

Numerical results

The simulated training data consists of a record $\{y_k\}$ of 4000 samples generated with Θ_0. We denote by θ_0 the AR(2) model identified by the least squares method. We also arbitrarily selected 5 other models which we denote by $\theta_1, \ldots, \theta_5$. The values of these θ are listed in table 1 and their corresponding poles are depicted in figure 2. Notice that models $\theta_3, \theta_4, \theta_5$ are unstable. In table 1 the Euclidian distances of the nominal models to θ_0 are also shown. We should have used a spectral distance to measure the model differences, but as far as we know, no spectral distance is well defined for unstable AR models.

Results of the test on signals generated using Θ_o are shown in table 2. Each row of the table represents the tests on a record $\{y_k\}$ of 1000 samples, with different nominal models. The last row shows the mean of the tests for each nominal model. Note that these results agree with the χ^2 behavior expected from Theorem 1. Similarly we generate 10 records $\{y_k\}$ of 1000 samples using Θ_1, and table 3 contains the corresponding tests. These numerical results show that it is not difficult to discriminate signals generated with Θ_o from those with Θ_1 by comparing the test values to some properly chosen threshold.

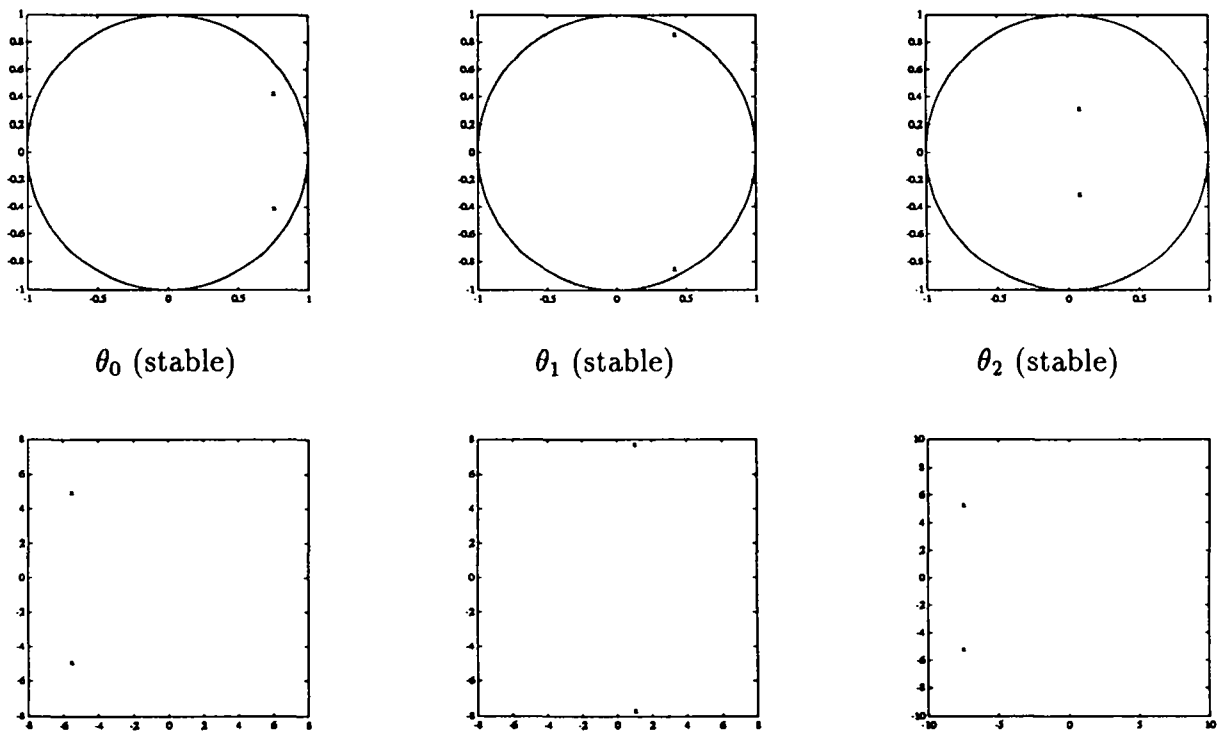
Discussion on the simulation results

The simulation results we presented in this subsection show that method 2 of our local approach can work with a reduced model structure and with some arbitrarily selected nominal

	θ_0	θ_1	θ_2	θ_3	θ_4	θ_5
a_1	-1.5092	-0.8339	-0.1729	11.0112	-2.0564	14.9847
a_2	0.7436	0.9059	0.1030	54.6210	59.8838	83.4328
Dist	0	0.6945	1.4820	55.3131	59.1427	84.3182

θ_0 is identified on the training data, the others are arbitrarily selected.
Dist is the Euclidian distance of the model parameters to θ_0.

Table 1: Nominal models



θ_0 (stable) θ_1 (stable) θ_2 (stable)

θ_3 (unstable) θ_4 (unstable) θ_5 (unstable)

Figure 2: Poles of the nominal models, c.f. table 1.

test number	θ_0	θ_1	θ_2	θ_3	θ_4	θ_5
1	2.6833	2.6624	2.8048	2.0019	2.2529	2.0127
2	1.0850	1.0773	0.9534	1.0107	1.1479	1.0209
3	7.5541	7.5142	7.7575	7.2582	7.3094	7.2550
4	1.6879	1.7061	1.6930	1.7690	1.8164	1.7708
5	4.3015	4.5761	4.6175	2.4641	3.2116	2.4910
6	4.4502	4.3049	4.4155	4.5101	4.5324	4.5124
7	0.4788	0.4789	0.4864	0.4626	0.4643	0.4629
8	7.4200	7.6835	7.8805	4.1991	5.6891	4.2620
9	3.2310	3.5769	3.7692	1.7660	2.2989	1.7786
10	2.3350	2.6546	2.9302	1.0360	1.4202	1.0413
mean	3.5227	3.6235	3.7308	2.6478	3.0143	2.6608

Table 2: Tests on the signals generated with Θ_o

test number	θ_0	θ_1	θ_2	θ_3	θ_4	θ_5
1	286.7550	287.4121	287.8741	195.2101	248.1481	197.9013
2	319.3328	316.4148	321.4840	243.7786	286.8403	245.9753
3	257.8440	254.5922	257.9128	190.0052	229.2414	192.0294
4	245.2815	245.9239	247.3391	176.6232	216.3758	178.6465
5	226.5726	224.7039	227.0450	169.4316	203.0966	171.1733
6	144.4646	147.0293	145.8105	93.8416	123.2575	95.3231
7	315.5679	312.1756	314.6995	227.4374	279.2217	230.1314
8	175.7328	174.4717	174.4225	124.8941	155.8752	126.5333
9	199.1021	195.8453	196.9424	147.9792	179.4706	149.6778
10	272.3833	271.8079	272.7751	203.6063	246.0436	205.8079
mean	244.3037	243.0377	244.6305	177.2807	216.7571	179.3199

Table 3: Tests on the signals generated with Θ_1

models. This illustrates the robustness of our approach with respect to model reductions. On the other hand, we see that the estimation of the basic statistics bias can actually replace the nominal model identification, which was considered as a necessary stage for the application of the local approach for change detection as presented in previous papers on this subject. Remark that even unstable models can be used as the nominal model. Though the selected nominal model can be quite far from the identified one, the numerical results show that by using the well identified nominal model θ_0 or by selecting models close to θ_0 we get higher test values under $\mathcal{H}_1$ than by selecting models far from θ_0, therefore the test is more efficient with "better" models. Note that when the nominal model θ_o is well identified (using an identification procedure associated with the basic statistics used for detection, as for the identification of θ_0 in this example), the bias h_o is zero, in this case method 2 is equivalent to method 1. This suggests us that method 1 should be used whenever possible, which is in agreement with the theoretical performance analysis of subsection 2.4.

3.2 Application to the monitoring of a gas turbine

We present in this subsection the application of our local approach to the monitoring of a gas turbine. Gas turbines are widely used as industrial motors, their typical uses are electrical power generators and aircraft engines. Usually a gas turbine system is mainly composed of a compressor, a set of combustion chambers and an expansion turbine. The purpose of our study is to monitor the sub-system composed of the combustion chambers and the expansion turbine[3]. First we present a simplified mathematical model of the system, then we describe the implementation of our local approach for change detection, and finally, we present the results on simulated data.

3.2.1 Simplified mathematical model

We are interested in monitoring the combustion chambers (there are 10 chambers in our system) and the expansion turbine. In the combustion chambers the fuel is mixed with compressed air and is then burned to produce a gas of high temperature and high pressure. This burned gas is then guided into the expansion turbine and drives the rotor by means of expansion.

The gas at the entry of the expansion turbine has a temperature profile generated by the 10 combustion chambers. The mean temperature of this profile T_e is observed with low accuracy[4], and we observe as well the temperature profile at the exhaust of the expansion turbine with 18 thermocouples. The complete modeling of such a gas turbine system can be found in [16], here we give a simplified model of the combustion chambers and the expansion turbine. After some simplifications the temperatures measured by the 18 thermocouples y_j

[3]This study is part of the TAG project of European Gas Turbine SA and Alcatel-Alsthom-Recherche in which IRISA participates

[4]in practice this is done indirectly by a complicated procedure.

can be expressed in the following way:

$$y_j = \frac{T_s}{T_e}\left(T_e + \sum_{i=0}^{9} a_i g(\phi_j^t - \phi_i^c - d)\right) + b\cos(\phi_j^t - \phi^s) + w_j, \quad j = 0, \ldots, 17 \tag{38}$$

$$d = D(T_e, T_s, N) + d_0 \tag{39}$$

where the notations are explained as follows. We use a zero mean function $g(\cdot)$ to express the shape of the temperature profile generated by each combustion chamber and a_i is the scaling factor of the i-th chamber. ϕ_i^c is the angular coordinate of the i-th combustion chamber, and ϕ_j^t is the angular coordinate of the j-th thermocouple. T_s denotes the mean temperature at the exhaust of the expansion turbine, so T_s/T_e expresses the fall of temperature during the expansion. Besides this fall of temperature, the gas rotates around the axis of the turbine during its expansion. We call this rotation angle *phase shift* and denote it by d. Using prior knowledge, we can compute the phase shift in an approximate way, and this computation is expressed by a known function $D(T_e, T_s, N)$ where N is the rotation speed of the turbine. In view of the approximate feature of $D(T_e, T_s, N)$, we use an additional parameter d_0 to compensate for its error. The term $b\cos(\phi_j^t - \phi^s)$ takes into account the geometric dissymmetry of the expansion turbine where ϕ^s depends on the direction of the exhauster and b gives a measurement of this dissymmetry. Finally w_j's are noises introduced due to modeling and measurement errors.

To summarize, the parameter vector of the model is written as:

$$\theta = (a_0, \ldots, a_9, b, d_0)^T$$

and the available observations are:

$$X = (T_e, T_s, N, y_0, \ldots, y_{17})^T \tag{40}$$

where T_s is a simple average of the 18 thermocouples. Since (38) is a static system and we assume the noise to be white, sampled X is an i.i.d. sequence. We shall monitor changes in the process using records of X which we denote by $\{X_k\}$.

An identification procedure was proposed in [16] to identify θ using a record $\{X_k\}$. It consists in attempting to minimize the cost function:

$$C(\theta) = \mathbf{E}\{c(\theta, X_k)\} \tag{41}$$

where

$$c(\theta, X_k) = \frac{1}{2}\sum_{j=0}^{17}(F_j(\theta, k) - y_j(k))^2 \tag{42}$$

and

$$F_j(\theta, k) \triangleq \frac{T_s(k)}{T_e}\left(T_e(k) + \sum_{i=0}^{9} a_i g(\phi_j^t - \phi_i^c - D(T_e(k), T_s(k), N(k)) - d_0)\right) + b\cos(\phi_j^t - \phi^s) \tag{43}$$

(notice that $y_j(k)$'s are measurements given by the thermocouples and $F_j(\theta, k)$'s are predictions given by the model). Note that a stochastic gradient algorithm of Robbins-Monro type would be of the form:

$$\theta_k = \theta_{k-1} + \gamma_k H(\theta_{k-1}, X_k) \tag{44}$$

with

$$H(\theta, X_k) = -\frac{\partial c(\theta, X_k)}{\partial \theta} = -\sum_{j=0}^{17} \left((F_j(\theta, k) - y_j(k)) \frac{\partial F_j(\theta, k)}{\partial \theta} \right) \tag{45}$$

Unfortunately, we were unable to handle this recursive algorithm in order to make it stable and convergent. The reasons for this are the nonlinearity of the model, and the presence of noise at both the output and *the input* of the considered system. So the basic idea of the identification procedure proposed in [16] was to separate d_0 from the other parameters during the identification and this results in an off-line identification procedure. In this way we manage to give a stable estimation of θ, but this estimation is known to be biased because of the noisy input [13]. The bias is essentially caused by the noises at the input of the identified process, that is, the measurement noises of T_e and N. Although a more complicated identification procedure may possibly remove this bias, we preferred to keep our relatively simple identification procedure to identify the nominal model of the system, because the purpose of this identification is to monitor changes in the process, and as we previously discussed, the local approach for change detection can be implemented using poor nominal models.

3.2.2 Implementation of the on-line detection

The system is simulated using equations (38)–(39), therefore the "true parameter" Θ_* of the system in this case is the simulation parameter:

$$\Theta_* = (a_0, \ldots, a_9, b, d_0)^T.$$

Denote by θ_o the nominal model identified via the previously discussed method. Because the identification of θ_o is not recursive, we cannot derive a basic statistics associated with that identification procedure. So we simply choose $H(\theta_o, X_k)$ defined by equation (45) for the basic statistics. Note that the X auxiliary variable is identical to the observation vector X in this case. Because the nominal model θ_o is not identified by an identification procedure associated with the basic statistics $H(\theta_o, X_k)$, method 1 presented in the previous section cannot be applied, so we have to apply method 2 which involves the modification of the basic statistics by removing from $H(\theta_o, X_k)$ its mean h_o as defined in (27). The estimations of h_o and R_o from $H(\theta_o, X_k)$ follow the same formulae as for the AR case presented in the previous subsection, except that the first term in estimator (35) suffices for estimating R_o in this case, because $H(\theta_o, X_k)$ is an i.i.d. sequence that in turn is the consequence of the static feature and noise nature of the system model (38).

We still denote the corrected basic statistics by

$$Z_k = H(\theta_o, X_k) - \hat{h}_o,$$

then the on-line detection procedure is summarized as follows.

can be expressed in the following way:

$$y_j = \frac{T_s}{T_e}\left(T_e + \sum_{i=0}^{9} a_i g(\phi_j^t - \phi_i^c - d)\right) + b\cos(\phi_j^t - \phi^s) + w_j, \quad j = 0, \ldots, 17 \tag{38}$$

$$d = D(T_e, T_s, N) + d_0 \tag{39}$$

where the notations are explained as follows. We use a zero mean function $g(\cdot)$ to express the shape of the temperature profile generated by each combustion chamber and a_i is the scaling factor of the i-th chamber. ϕ_i^c is the angular coordinate of the i-th combustion chamber, and ϕ_j^t is the angular coordinate of the j-th thermocouple. T_s denotes the mean temperature at the exhaust of the expansion turbine, so T_s/T_e expresses the fall of temperature during the expansion. Besides this fall of temperature, the gas rotates around the axis of the turbine during its expansion. We call this rotation angle *phase shift* and denote it by d. Using prior knowledge, we can compute the phase shift in an approximate way, and this computation is expressed by a known function $D(T_e, T_s, N)$ where N is the rotation speed of the turbine. In view of the approximate feature of $D(T_e, T_s, N)$, we use an additional parameter d_0 to compensate for its error. The term $b\cos(\phi_j^t - \phi^s)$ takes into account the geometric dissymmetry of the expansion turbine where ϕ^s depends on the direction of the exhauster and b gives a measurement of this dissymmetry. Finally w_j's are noises introduced due to modeling and measurement errors.

To summarize, the parameter vector of the model is written as:

$$\theta = (a_0, \ldots, a_9, b, d_0)^T$$

and the available observations are:

$$X = (T_e, T_s, N, y_0, \ldots, y_{17})^T \tag{40}$$

where T_s is a simple average of the 18 thermocouples. Since (38) is a static system and we assume the noise to be white, sampled X is an i.i.d. sequence. We shall monitor changes in the process using records of X which we denote by $\{X_k\}$.

An identification procedure was proposed in [16] to identify θ using a record $\{X_k\}$. It consists in attempting to minimize the cost function:

$$C(\theta) = \mathbf{E}\{c(\theta, X_k)\} \tag{41}$$

where

$$c(\theta, X_k) = \frac{1}{2}\sum_{j=0}^{17}(F_j(\theta, k) - y_j(k))^2 \tag{42}$$

and

$$F_j(\theta, k) \triangleq \frac{T_s(k)}{T_e}\left(T_e(k) + \sum_{i=0}^{9} a_i g(\phi_j^t - \phi_i^c - D(T_e(k), T_s(k), N(k)) - d_0)\right) + b\cos(\phi_j^t - \phi^s) \tag{43}$$

(notice that $y_j(k)$'s are measurements given by the thermocouples and $F_j(\theta, k)$'s are predictions given by the model). Note that a stochastic gradient algorithm of Robbins-Monro type would be of the form:

$$\theta_k = \theta_{k-1} + \gamma_k H(\theta_{k-1}, X_k) \tag{44}$$

with

$$H(\theta, X_k) = -\frac{\partial c(\theta, X_k)}{\partial \theta} = -\sum_{j=0}^{17} \left((F_j(\theta, k) - y_j(k)) \frac{\partial F_j(\theta, k)}{\partial \theta} \right) \tag{45}$$

Unfortunately, we were unable to handle this recursive algorithm in order to make it stable and convergent. The reasons for this are the nonlinearity of the model, and the presence of noise at both the output and *the input* of the considered system. So the basic idea of the identification procedure proposed in [16] was to separate d_0 from the other parameters during the identification and this results in an off-line identification procedure. In this way we manage to give a stable estimation of θ, but this estimation is known to be biased because of the noisy input [13]. The bias is essentially caused by the noises at the input of the identified process, that is, the measurement noises of T_e and N. Although a more complicated identification procedure may possibly remove this bias, we preferred to keep our relatively simple identification procedure to identify the nominal model of the system, because the purpose of this identification is to monitor changes in the process, and as we previously discussed, the local approach for change detection can be implemented using poor nominal models.

3.2.2 Implementation of the on-line detection

The system is simulated using equations (38)-(39), therefore the "true parameter" Θ_* of the system in this case is the simulation parameter:

$$\Theta_* = (a_0, \ldots, a_9, b, d_0)^T.$$

Denote by θ_o the nominal model identified via the previously discussed method. Because the identification of θ_o is not recursive, we cannot derive a basic statistics associated with that identification procedure. So we simply choose $H(\theta_o, X_k)$ defined by equation (45) for the basic statistics. Note that the X auxiliary variable is identical to the observation vector X in this case. Because the nominal model θ_o is not identified by an identification procedure associated with the basic statistics $H(\theta_o, X_k)$, method 1 presented in the previous section cannot be applied, so we have to apply method 2 which involves the modification of the basic statistics by removing from $H(\theta_o, X_k)$ its mean h_o as defined in (27). The estimations of h_o and R_o from $H(\theta_o, X_k)$ follow the same formulae as for the AR case presented in the previous subsection, except that the first term in estimator (35) suffices for estimating R_o in this case, because $H(\theta_o, X_k)$ is an i.i.d. sequence that in turn is the consequence of the static feature and noise nature of the system model (38).

We still denote the corrected basic statistics by

$$Z_k = H(\theta_o, X_k) - \hat{h}_o,$$

then the on-line detection procedure is summarized as follows.

On-line detection procedure :

1. At time n, we have a record of observations

$$\{X_k : k = 1, \ldots, n\}$$

2. compute for $r \leq n$

$$D_r^n = (n - r + 1)^{-\frac{1}{2}} \sum_{k=r}^{n} Z_k \tag{46}$$

$$S_r^n = (D_r^n)^T \hat{R}_o^{-1} D_r^n. \tag{47}$$

3. the decision rule with a chosen threshold λ is:

$$\max_{r \in M} S_r^n \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda.$$

4. When a change is detected ($\mathcal{H}_1$ has been decided), the detection time is:

$$\nu = \min\{n : \max_{r \in M} S_r^n \geq \lambda\}$$

and the estimated change time is given by:

$$\hat{r} = \arg\max_{r \in M} S_r^\nu.$$

In the last three formulas, M is a set of time instants at which a possible change is searched up to time n. For efficiency, in the experiments reported below, r is taken within a moving window $M \triangleq \{(n - n_1), \ldots, (n - n_0)\}$, where we have chosen $n_0 = 50$ and $n_1 = 500$. With respect to the maximal window $M = \{1, \ldots, n\}$, this choice does have a negative influence on the delay for detection and on the estimated change time as will be shown next, but a positive influence on the false alarm rate.

3.2.3 Results of simulation

We use the model described in equations (38)–(39) to simulate the measurement data. In table 4 we show the simulation parameter Θ_* used to generate data under $\mathcal{H}_0$ and the nominal model parameter θ_o obtained by identification over a record of 2000 observation samples. We have simulated the observations X_k with additive white noises the standard deviations of which are 10% of the ranges of the simulated signals. The nominal model as shown in table 4 is obviously biased and is used in the previously described implementation.

Three on-line test simulations are shown in figures 3, 4 and 5. In each of the three simulations we simulated a different change in the parameter Θ_*. See the comments in figures 3, 4 and 5. In all these simulations we have chosen the test threshold $\lambda = 40$.

	a_0	a_1	a_2	a_3	a_4	a_5
Θ_*	99.55	92.01	92.34	108.61	99.43	105.13
θ_o	83.17	78.18	80.74	95.59	83.79	88.14

	a_6	a_7	a_8	a_9	b	d_0
Θ_*	103.96	107.56	104.00	86.58	60.0	0.1
θ_o	88.90	94.98	90.44	71.85	59.71	0.1029

Table 4: Simulation parameter under $\mathcal{H}_0$ and nominal model

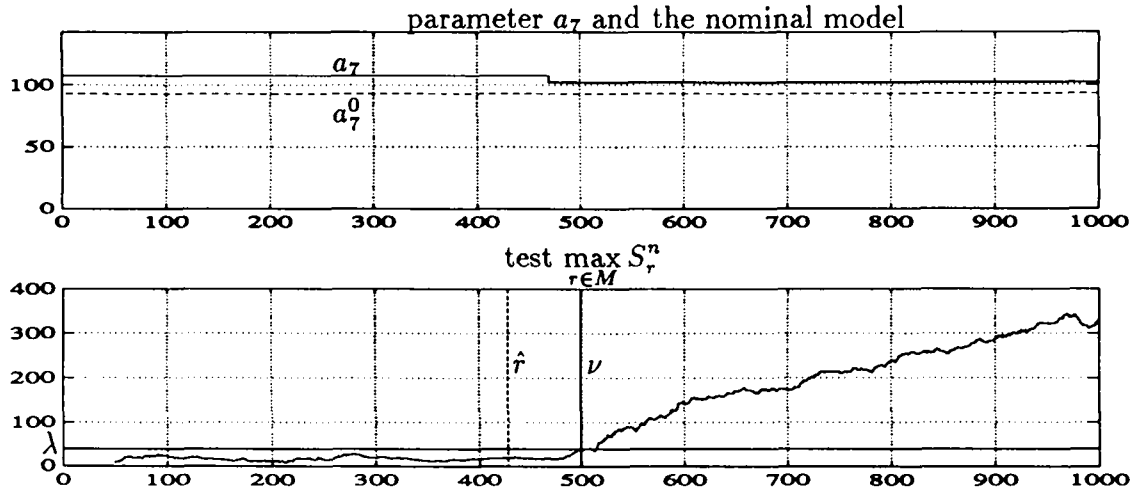
According to the theoretical issues in [6, 7], the test value under $\mathcal{H}_0$ should be about 12, the number of parameters in the model, independent of the number of samples on which the test is made; on the other hand, the test value under $\mathcal{H}_1$ is higher than under $\mathcal{H}_0$, and the larger the number of samples is, the higher the test value is. This is an indication for the choice of λ, but the independency of the test value under $\mathcal{H}_0$ on the number of test samples is under the assumption that the nominal model is perfectly identified for method 1, or equivalently, the bias of the basic statistics is perfectly estimated for method 2. In practice this "perfectness" is never guaranteed, consequently the threshold choice should be a bit higher, and slightly depends on the length of the used moving window M. Furthermore, the choice of λ is a tradeoff between the delay for detection and the false alarm rate.

These results show the efficiency of the local approach for detecting small changes, in particular its robustness with respect to the bias of nominal model identification. After the theoretical discussions in the previous sections, it should not be a surprise to see that our approach successfully tests the change of a parameter which becomes, after the change, closer to the nominal model in use, as what is shown by the first example in figure 3. This confirms that we do not necessarily need to know the true parameter of the system in order to detect change in it, instead we can use a nominal model with possibly different parameters. This robustness is important in practice, because models used to monitor systems are often simplified and identification procedures are never perfect.

4 Conclusion

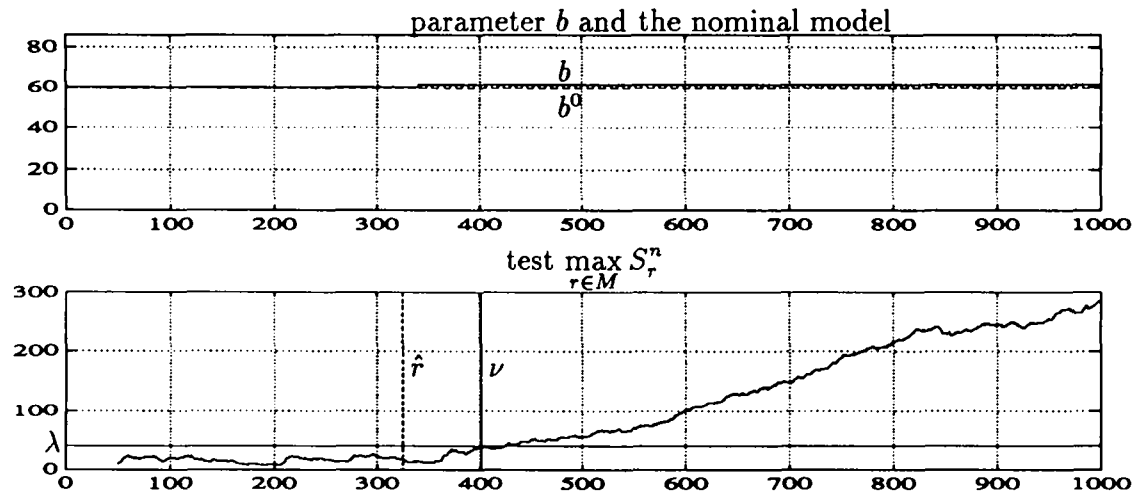
Techniques for early warning of slight changes in systems and plants are useful for condition-based maintenance. In this paper we present an approach for this problem. This approach is based on the so-called "asymptotic local" approach for change detection previously introduced by the same authors. We have presented in this paper a considerable generalization of the original method as presented in [6, 7]. The features of our generalized approach are the following :

- The *models* used for monitoring can be different from the *systems* to which monitoring is applied; in particular, model reduction, can be enforced.



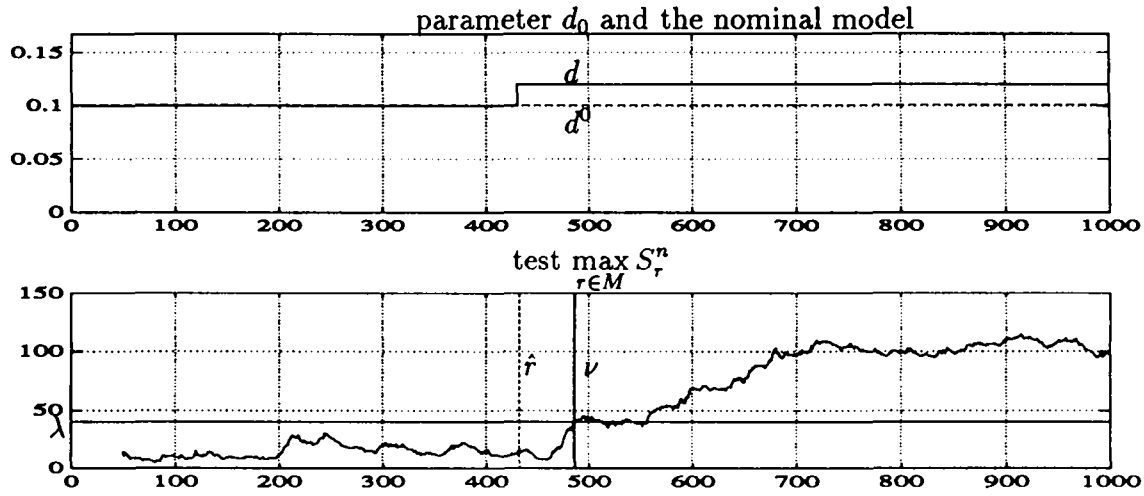
a_7 changes from 107.56 to 102.2 at time $r = 470$.
Detection time $\nu = 499$, delay for detection $\nu - r = 29$, estimated change time $\hat{r} = 428$.

Figure 3: detect a change of a_7



b changes from 60.0 to 61.20 at time $r = 340$.
Detection time $\nu = 401$, delay for detection $\nu - r = 61$, estimated change time $\hat{r} = 325$.

Figure 4: detect a change of b



d_0 changes from 0.1 to 0.12 at time $r = 430$.
Detection time $\nu = 486$, delay for detection $\nu - r = 56$, estimated change time $\hat{r} = 432$.

Figure 5: detect a change of d_0

- The central tool in our approach is the *basic statistics* $H(\theta, X_k)$, where θ is the adjustable parameter characterizing a particular model, and X_k summarizes actual measurements from the system. Typically, the basic statistics originates from some recursive identification procedure of the form :

$$\theta_k = \theta_{k-1} + \gamma_k H(\theta_{k-1}, X_k) \tag{48}$$

  but this is not mandatory.

- Thanks to some central limit theorem, we have shown that monitoring the system reduces to testing for changes in the mean of the sequence :

$$Z_k \triangleq H(\theta_o, X_k) \tag{49}$$

  as if Z_k was i.i.d. and Gaussian. In (49), θ_o is a *nominal model* which can be either (*for Method 1*) identified using the algorithm (48), or (*for Method 2*) arbitrarily chosen; then the bias of Z_k under $\mathcal{H}_o$ must be Monte Carlo estimated and removed.

We have reported simulation results concerning AR models and preliminary results from a real application in the area of condition based maintenance.

Appendix

A Mathematical results for on-line algorithms

We now consider Problem 2 stated in section 2.3. The Cusum (19) is modified as follows

$$\text{for } m \leq N \ : \quad D_{N,m}(\theta) \triangleq \frac{1}{\sqrt{N}} \sum_{k=1}^{m} H(\theta, X_k) \tag{50}$$

and we define the continuous time stochastic process $(D_{N,t})_{0 \leq t \leq 1}$ as the piecewise linear interpolation of the discrete sequence $D_{N,m}$ at the points $t = m/N$. Then theorem 1 rewrites as follows:

Theorem 2 *Under Conditions 1, 2, 3, and the hypothesis : there exists some unknown* $\tau \in (0,1)$ *such that*

$$\begin{aligned} \mathcal{F}(\Theta_*) &= \theta_o \quad \text{for} \quad 1 \leq k < \tau N \\ \mathcal{F}(\Theta_*) &= \theta_o + \frac{\tilde{\theta}}{\sqrt{N}} \quad \text{for} \quad \tau N \leq k \leq N \end{aligned}$$

when $N \to \infty$, *the Cusum process* $(D_{N,t}(\theta_o))_{0 \leq t \leq 1}$ *converges weakly to the following Brownian motion with changing drift :*

$$dD_t = -\mathbf{1}_{\{t \leq \tau\}} \, h'(\Theta_*) \cdot \tilde{\theta} \, dt + R^{1/2}(\Theta_*) \, dW_t \tag{51}$$

where $h'(\Theta_*)$ *and* $R(\Theta_*)$ *are defined in (23) and (24) respectively, and* W_t *denotes a standard Brownian motion.*

The same comments as for Theorem 1 apply here. To get the proof of this theorem, just perform the same modification to the proof of the appendix of [6] as we have done for proving Theorem 1, see also Chapter 5 of Part I of [7]. This justifies the reduction of on-line change detection problem to the Gaussian equivalent one similar to that stated in Problem 3.

References

[1] R. ANDRÉ-OBRECHT, "A new statistical approach for automatic segmentation of continuous speech signals", *IEEE Trans. ASSP*, vol. ASSP-36, no. 1 pp. 29-40, Jan.1988.

[2] M. BASSEVILLE, "Detecting changes in signals and systems - A survey", *Automatica*, vol. 24, no 3, pp. 309-326, May 1988.

[3] M. BASSEVILLE, A. BENVENISTE, B. GACH-DEVAUCHELLE, M. GOURSAT, D. BONNECASE, P. DOREY, M. PREVOSTO, M. OLAGNON, "Damage monitoring in vibration mechanics: issues in diagnostics and predictive maintenance", *IRISA Res. Report*, submitted for publication, 1992.

[4] M.BASSEVILLE, A.BENVENISTE, G.MOUSTAKIDES, A.ROUGÉE: "Optimal sensor location for detecting changes in dynamical behavior", *IEEE Trans. on Automatic Control*, vol.AC-32 n°12, 1067-1075, Dec. 1987.

[5] M.BASSEVILLE, I.V.NIKIFOROV, *Detection of Abrupt Changes - Theory and Applications*, Prentice Hall Information and System Sciences Series, to appear.

[6] A.BENVENISTE, M.BASSEVILLE, G.MOUSTAKIDES: "The asymptotic local approach to change detection and model validation". *IEEE Trans. on Automatic Control*, vol. AC-32, no.7, pp.583-592. July 1987.

[7] A. BENVENISTE, M. METIVIER, P. PRIOURET, *Adaptive Algorithms and Stochastic Approximations*, Applications of Mathematics, Vol. 22, Springer Verlag, 1990.

[8] B. DEVAUCHELLE-GACH, *"Diagnostic mécanique des fatigues sur les structures soumises à des vibrations en ambiance de travail"*, Thesis, Univ. Paris IX Dauphine, Dec. 1991, (in French).

[9] J. DIEUDONNE, *Fondements de l'Analyse Moderne*, Gauthier-Villars, Paris, 1968.

[10] I.V. NIKIFOROV AND I.N. TIKHONOV "Application of change detection theory to seismic signal processing" in *Detection of Abrupt Changes in Signals and Dynamical Systems*, M. Basseville and A. Benveniste, Eds, Lecture Notes in Control and Information Sciences, LNCIS 77, Springer-Verlag, New York, pp.355-373, 1986.

[11] R.J. PATTON, P.M. FRANK AND R.N. CLARK (EDS), *Fault Diagnosis in Dynamic Systems : Theory and Application*, Prentice Hall International (UK), 1989.

[12] G. G. ROUSSAS, *Contiguity of Probability Measures, some Applications in Statistics.* Cambridge University Press, 1972.

[13] T.SÖDERSTRÖM, "Identification of stochastic linear systems in presence of input noise", *Automatica*, vol.17, pp.713-725, 1981.

[14] A.S. WILLSKY, "A survey of design methods for failure detection in dynamic systems", *Automatica*, vol. 12, pp. 601-611, 1976.

[15] A.S. WILLSKY, "Detection of abrupt changes in dynamic systems", in *Detection of Abrupt Changes in Signals and Dynamical Systems*, M. Basseville and A. Benveniste, Eds, Lecture Notes in Control and Information Sciences, LNCIS 77, Springer Verlag, 27-49, 1986.

[16] Q. ZHANG, *"Contribution à la surveillance de procédés industriels"*, Thesis, Univ. Rennes I, Dec. 1991, (in French).

LISTE DES DERNIERES PUBLICATIONS INTERNES PARUES A L'IRISA

PI 661 REACHABILITY ANALYSIS ON DISTRIBUTED EXECUTIONS
Claire DIEHL, Claude JARD, Jean-Xavier RAMPON
Juin 1992, 18 pages.

PI 662 RECONSTRUCTION 3D DE PRIMITIVES GEOMETRIQUES PAR VISION ACTIVE
Samia BOUKIR, François CHAUMETTE
Juin 1992, 40 pages.

PI 663 FILTRES SEMANTIQUES EN CALCUL PROPOSITIONNEL
Raymond ROLLAND
Juin 1992, 22 pages.

PI 664 REGION-BASED TRACKING IN AN IMAGE SEQUENCE
François MEYER, Patrick BOUTHEMY
Juin 1992, 50 pages.

PI 665 CORRECTNESS OF AUTOMATED DISTRIBUTION OF SEQUENTIAL PROGRAMS
Cyrille BARREAU, Benoît CAILLAUD, Claude JARD, René THORAVAL
Juin 1992, 32 pages.

PI 666 AGREGATION FAIBLE DES PROCESSUS DE MARKOV ABSORBANTS
James LEDOUX, Gerardo RUBINO, Bruno SERICOLA
Juillet 1992, 30 pages.

PI 667 MODELES D'EVALUATION DE LA FIABILITE DU LOGICIEL ET TECHNIQUES DE VALIDATION DE SYSTEMES DE PREDICTION : ETUDE BIBLIOGRAPHIQUE
James LEDOUX
Juillet 1992, 76 pages.

PI 668 TWO COMPLEMENTARY NOTES ON SKEWED-ASSOCIATIVE CACHES
André SEZNEC
Juillet 1992, 10 pages.

PI 669 PARALLELISATION D'UN ALGORITHME DE DETECTION DE MOUVEMENT SUR UNE ARCHITECTURE MIMD
Fabrice HEITZ, Sergui JUFRESA, Etienne MEMIN, Thierry PRIOL
Juillet 1992, 34 pages.

PI 670 UN RESEAU SYSTOLIQUE INTEGRE POUR LA CORRECTION DE FAUTES DE FRAPPE
Dominique LAVENIER
Juillet 1992, 120 pages.

PI 671 EARLY WARNING OF SLIGHT CHANGES IN SYSTEMS AND PLANTS WITH APPLICATION TO CONDITION BASED MAINTENANCE
Qinghua ZHANG, Michèle BASSEVILLE, Albert BENVENISTE
Juillet 1992, 32 pages.

PI 672 ORDRES REPRESENTABLES PAR DES TRANSLATIONS DE SEGMENTS DANS LE PLAN
Vincent BOUCHITTE, Roland JEGOU, Jean-Xavier RAMPON
Juillet 1992, 8 pages.

ISSN 0249 - 6399